Project Planning Phase

**Project Planning Template (Product Backlog, Sprint Planning, Stories, Story points)**

|  |  |
| --- | --- |
| **Field** | **Details** |
| **Date** | 28 June 2025 |
| **Team ID** | LTVIP2025TMID45560 |
| **Project Name** | Revolutionizing Liver Care: Predicting Liver Cirrhosis using Advanced Machine Learning  Techniques |
| **Team Members** | MASABATTULA DIVYA (S201086), KOTA HEMALATHA (S200381) |
| **Maximum**  **Marks** | 5 Marks |
|  |  |
|  |  |

**Product Backlog, Sprint Schedule, and Estimation (4 Marks)**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Sprint** | **Functional Requirement**  **(Epic)** | **User Story**  **Number** | **User Story / Task** | **Story Points** | **Priority** | **Team Members** |
| **Sprint- 1** | **Data Collection & Setup** | USN-1 | As a data scientist, I need to collect and acquire liver disease dataset with clinical parameters  for model training | 3 | High | Divya, Hemalatha |
| Sprint- 1 |  | USN-2 | As a developer, I need to set up the development environment with Python, pandas, scikit-learn,  and other ML libraries | 2 | High | Divya |
| Sprint- 1 |  | USN-3 | As a data analyst, I need to perform initial data exploration to understand dataset structure  and features | 2 | High | Hemalatha |
| **Sprint- 2** | **Data Preprocessing** | USN-4 | As a data scientist, I need to  handle missing values and outliers in the dataset | 3 | High | Hemalatha |
| Sprint- 2 |  | USN-5 | As a data scientist, I need to encode categorical variables and  normalize numerical features | 2 | High | Divya |
| Sprint- 2 |  | USN-6 | As a data analyst, I need to  perform feature selection and correlation analysis | 3 | Medium | Divya, Hemalatha |
| **Sprint- 3** | **Model Development** | USN-7 | As a ML engineer, I need to implement Logistic Regression model for liver cirrhosis  prediction | 2 | High | Divya |
| Sprint- 3 |  | USN-8 | As a ML engineer, I need to  implement KNN and SVM models for comparison | 3 | High | Hemalatha |
| Sprint- 3 |  | USN-9 | As a ML engineer, I need to implement Random Forest and  XGBoost models | 3 | High | Divya |
| Sprint- 3 |  | USN-10 | As a developer, I need to split data into training and testing sets | 1 | High | Hemalatha |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Sprint** | **Functional Requirement (Epic)** | **User Story Number** | **User Story / Task** (80:20 ratio) | **Story Points** | **Priority** | **Team Members** |
| **Sprint- 4** | **Model Evaluation** | USN-11 | As a ML engineer, I need to evaluate all models using accuracy, precision, recall, and F1-  score | 3 | High | Divya, Hemalatha |
| Sprint- 4 |  | USN-12 | As a data scientist, I need to perform hyperparameter tuning  for best model performance | 4 | High | Divya |
| Sprint- 4 |  | USN-13 | As a analyst, I need to create confusion matrices and  performance comparison tables | 2 | Medium | Hemalatha |
| **Sprint- 5** | **Web Application Development** | USN-14 | As a developer, I need to create  Flask backend application for model deployment | 4 | High | Divya |
| Sprint- 5 |  | USN-15 | As a frontend developer, I need  to create HTML forms for user input of clinical parameters | 3 | High | Hemalatha |
| Sprint- 5 |  | USN-16 | As a developer, I need to integrate trained ML model with  Flask application | 3 | High | Divya |
| **Sprint- 6** | **Testing & Documentation** | USN-17 | As a developer, I need to test the  web application with sample data inputs | 2 | High | Hemalatha |
| Sprint- 6 |  | USN-18 | As a team member, I need to  create comprehensive project documentation and thesis | 4 | High | Divya, Hemalatha |
| Sprint- 6 |  | USN-19 | As a developer, I need to prepare deployment demo and  screenshots | 2 | Medium | Divya |
|  |  |  |  |  |  |  |

**Project Tracker, Velocity & Burndown Chart (4 Marks)**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Sprint** | **Total Story**  **Points** | **Duration** | **Sprint Start**  **Date** | **Sprint End Date (Planned)** | **Story Points Completed (as on Planned End Date)** | **Sprint Release Date (Actual)** |
| **Sprint-**  **1** | 7 | 3 Days | 15 June  2025 | 17 June 2025 | 7 | 17 June 2025 |
| **Sprint-**  **2** | 8 | 4 Days | 18 June  2025 | 21 June 2025 | 8 | 21 June 2025 |
| **Sprint-**  **3** | 9 | 4 Days | 22 June  2025 | 25 June 2025 | 9 | 25 June 2025 |
| **Sprint-**  **4** | 9 | 3 Days | 26 June  2025 | 28 June 2025 | 9 | 28 June 2025 |
| **Sprint-**  **5** | 10 | 3 Days | 29 June  2025 | 1 July 2025 | - | - |
| **Sprint-**  **6** | 8 | 2 Days | 2 July  2025 | 3 July 2025 | - | - |
|  |  |  |  |  |  |  |

**Total Project Duration:** 19 Days **Total Story Points:** 51

**Velocity Calculation**

# Team Velocity Analysis:

**Sprint 1:** 7 points ÷ 3 days = 2.33 points/day

**Sprint 2:** 8 points ÷ 4 days = 2.0 points/day

**Sprint 3:** 9 points ÷ 4 days = 2.25 points/day

 **Sprint 4:** 9 points ÷ 3 days = 3.0 points/day

**Average Velocity:** (2.33 + 2.0 + 2.25 + 3.0) ÷ 4 = **2.39 points per day**

**Team Capacity:** 2 team members × 6 hours/day = 12 hours/day **Velocity per Hour:** 2.39 ÷ 12 = **0.199 story points per hour**

**Burndown Chart Analysis**

**Ideal vs Actual Progress:**

|  |  |  |
| --- | --- | --- |
| **Day** | **Remaining Story Points (Ideal)** | **Remaining Story Points (Actual)** |
| Day 0 | 51 | 51 |
| Day 3 | 44 | 44 (Sprint 1 Complete) |
| Day 7 | 36 | 36 (Sprint 2 Complete) |
| Day 11 | 27 | 27 (Sprint 3 Complete) |
| Day 14 | 18 | 18 (Sprint 4 Complete) |
| Day 17 | 8 | TBD (Sprint 5 in Progress) |
| Day 19 | 0 | TBD (Sprint 6 Target) |
|  |  |  |

**Risk Management & Mitigation**

|  |  |  |  |
| --- | --- | --- | --- |
| **Risk** | **Impact** | **Probability** | **Mitigation Strategy** |
| Dataset quality issues | High | Medium | Implement robust data validation and cleaning  procedures |
| Model performance below  expectations | High | Low | Use ensemble methods and hyperparameter tuning |
| Flask deployment issues | Medium | Low | Test deployment early and maintain backup plans |
| Time constraints | Medium | Medium | Prioritize high-value features and maintain MVP  scope |
|  |  |  |  |

**Definition of Done (DoD)**

# For Data Science Tasks:

 Code is tested and validated

 Performance metrics documented  Code is peer-reviewed

 Results are reproducible **For Development Tasks:**

 Feature is implemented and tested  Code follows coding standards

 Documentation is updated

 Integration testing completed **For Documentation Tasks:**

 Content is technically accurate  Formatting follows template  Peer review completed

 Final version approved

**Sprint Goals**

**Sprint 1:** Establish project foundation with data collection and environment setup **Sprint 2:** Clean and prepare data for machine learning model development **Sprint 3:** Develop and train multiple ML models for liver cirrhosis prediction **Sprint 4:** Evaluate and optimize model performance with comprehensive testing **Sprint 5:** Deploy best-performing model as a web application using Flask

**Sprint 6:** Complete testing, documentation, and project delivery

**Success Criteria**

1. **Technical:** Achieve >85% accuracy in liver cirrhosis prediction
2. **Functional:** Deploy working web application for real-time predictions
3. **Documentation:** Complete comprehensive thesis and technical documentation
4. **Timeline:** Deliver project within planned 19-day timeline
5. **Quality:** Pass all testing phases and peer reviews